**Data Release Planning Document**

Each data release will involve three planforms: Imperiia Scalar site, Imperiia Dataverse hub, and the DCRES website

1. Data files will be stored on Dataverse
2. A descriptive page of the dataset will be created on the Scalar site under the “Data Catalog” section.
3. If the dataset is related to a project other than Data Drive, then the data page will be attached to the project page on scalar as well as the project page on the Davis Center website (as Featured Content).
4. An event will be created on the Davis Center website. The dataset description will be copy-pasted to the event page. The Scalar data description page will be attached to the event page through a button link. The event page will be automatically attached to either the Imperiia page or a project page.
5. \*The transcript of the event will be published as a Basic Page and be attached to either the Imperiia page or a project page as Featured Content?
6. In this way, each data release will be advertised through the following channels: DCRES newsletter & social media (the event page); Imperiia social media (both event page + Scalar page); DCRES YouTube (a new series with event recordings)

**New description (Why Not):**

We constantly ask “why,” but... why not? We’re used to learning history through written text, oral narratives, and images, but why not numbers, lines, points, and polygons? We’re used to articles and monographs as “standard” forms of historical scholarship, but why not datasets, web maps, board games, and podcasts? We’re used to lengthy lectures, seminars, and presentations as ways of academic communications, but why not a casual, low-stress 15-minute chat?

In the “Why Not” event series by the Imperiia Project, we chat about the idiosyncrasies of our work in idiosyncratic ways: describing the Russian Empire in terms of numbers and shape; advocating data by highlighting its inconsistencies; introducing maps without showing them to you. Through this approach, we challenge the conventional ways of studying and communicating history and ask the question, "Why not?"

------

Our first chat is dedicated to a new data publication of the Imperiia Project: “Russians in the World.” Join the conversation with two data editors, Dr. Kelly O’Neill and Vivian Wei, to learn about how you may explore the social history of tsarist migration with passenger records from American ports.

More about the data (copy-pasting data description from the Scalar page):

## What it is

This is an open-access dataset intended for research and analysis. Its source is a dataset provided by the U. S. National Archives consisting of half a million passenger arrival records and ship manifests across six decades, from 1834 to 1897. Our edition consists of 7 related tables (csv), 2 spatial data files (shapefiles), and 2 metadata files (ReadMe; codebook). Together the files describe 527,394 passengers, 10,761 voyages, 780 ships, 705 occupations, 150 routes, and 78 ports.

## Why it matters

The original dataset contains passenger records with name, age, town of last residence, destination, and codes for sex, occupation, literacy, country of origin, transit and/or travel compartment. It also contains manifest records - think of them as voyage records - including ship name, arrival date, and arrival port.

As unique and vast as it is, the original data is not easy to use. The records are full of inconsistencies and ambiguities, and are not easily ingested by visualization and GIS tools. They are, like most historical sources, a wonderful example of messy data. Our work to generate an enhanced and usable edition fell into four buckets:

* We cleaned and tidied, documenting our process every step of the way.
* We reorganized the data according to a relational model.
* We added calculated fields as well as context fields drawn from other sources.
* We created vector data - port locations and route lines - to facilitate spatial analysis.

Fear not: inconsistencies and ambiguities still turn up everywhere, but we are relatively confident that with this new edition you will find them productive and stimulating rather than frustrating.

## Where to find it

The dataset is free and open to the public for non-commercial use. Click on the button below!

**LINK TO THE SCALAR PAGE**

**Add the link to the scalar page to the DCRES project page**

**Chat structure**

* **Very brief introduction**
* **Thematic takeaway**
* **Spatial (?) takeaway**
* **Methodological takeaway**
* **Personal takeaway**

---------

The event page will look like this:

**Title**: Data Release: Russians in the World

Maybe a new series: **Data Chat**? Something like that

Format: semi-structured, informal conversation about Imperiia’s datasets

Length: 15-20 mins

**Description:**

We are used to learning history through stories embedded in written text, oral narrations, and/or images, but how about studying history through numbers and shapes? **Data Chat** series introduces freely accessible, well-documented, utterly fascinating data about the world that was once the Russian Empire to a primarily English-speaking audience. All the data we discuss is produced and released by the Imperiia Project from historical sources. While our datasets cover various topics, most of them address change over time as well as change across space, making them especially attractive for people interested in spatial analysis and historical GIS. Each Data Chat is an informal conversation about various aspects of a specific dataset: what it contains, how it was produced, why it matters, and its possible applications.

By producing, publishing, and discussing historical data, we want to advance conversations about data-production as a form of scholarship and new ways of studying history.

Our first Data Chat is dedicated to a new data publication: “Russians in the World.” Join the conversation with two data editors, Dr. Kelly O’Neill and Vivian Wei, to learn about how you may explore the social history of tsarist migration with passenger records from American ports.

------